

周报

时间：2015-5-4 to 2015-5-10

汇报人：林明

本周完成：

1. 完善全局精简本体。
 - a) 丰富 LinkedIn 相关的本体的类和属性
2. TwitLogic 调研
 - a) 操作系统：只能在 linux、mac 上运行，windows 需要在虚拟机下运行，或者配合类 unix 兼容软件运行
 - b) 该工具并不是以 jar 包或者 API 调用的模式，可方便得供 java 代码调用，而是通过在本地起一个服务器，然后再运行一个客户端根据配置文件，不断查找保存相关内容，这样工作方式。和我们系统的初衷（在线即时查询）不符。
 - c) 工具只能更具输入的用户名，用户列表，关键字；获得相应用户、用户列表的推文，或与关键字相关的推文等信息。而目前我们希望从 twitter 拿到的数据是：由关键字→匹配出来的用户→的推文。难以用这个工具实现
 - d) 所以，twitter 的爬虫应该不考虑此工具
3. 后端的重构工作，由于目前各个在线源的后端都在重写，于是提前展开。按照之前的规划，将后端统一规划为：网页获取\API 调用、信息解析、本体化三大模块。但是实际代码过程中，发现网页获取\API 调用、信息解析的解耦合难度较大，目前没有什么比较好的想法，本体化模块以写出雏形，但是周五讨论后，发现还有些漏洞，会尽快修复。
4. 5 月 8 号知识图谱研讨会
 - a) 个人的体会是，报告的内容深浅相差较大。通俗的报告只讲了些知识图谱领域的基础概念和理论结构，这么我们已经学过；深入的报告，都在他们已有的项目上，讨论深层地的难而细的关键问题，我目前的水平未能跟上；另外有些项目的核心感觉应该不在知识图谱上，不过确实是用到了知识图谱。
 - b) 但是收获还是很多的，最重要的是见识到了一些新的概念、应用、工具。还有对目前业内的工作认识也被刷新，感觉工业界已经实现的功能其实还蛮厉害的。不过也只是一种模糊的概念和感觉。
5. TextDS 前后端的衔接
 - a) 后端这周的工具量较少，因为上周基本的接口和函数已大部分实现，主要还是前段先实现一定的显示功能，然后联合 debug

下周计划：

1. TextDS 前后端的连接。
2. CMU neil.调研。
3. 后端重构工作。